DOCUMENT RESUME

ED 470 748 CS 511 508

AUTHOR	Rabinowitz, Stanley; Wong, Judy; Filby, Nikola
TITLE	Understanding Young Readers: The Role of Early Literacy Assessment. Knowledge Brief.
INSTITUTION	WestEd, San Francisco, CA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	2002-00-00
NOTE	17p.
CONTRACT	ED-01-CO-0012
AVAILABLE FROM	WestEd, 730 Harrison St., San Francisco, CA 94107-1242 (Product no. KN-02-02, $8). Tel: 877-493-7833 (Toll Free); Tel: 415-565-3000; Web site: http://www.WestEd.org/. For full text: http://www.WestEd.org/online_pubs/kn-02-02.pdf.
PUB TYPE	Guides - Non-Classroom (055) -- Opinion Papers (120)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Alternative Assessment; *Beginning Reading; *Emergent Literacy; *Evaluation Methods; Instructional Effectiveness; *Measurement Techniques; Primary Education; *Reading Achievement; *Student Evaluation

ABSTRACT

With reading, as with all academic content and skills, effective instruction is informed by sound assessment. Teachers' knowledge about classroom-embedded reading assessment must continue to be developed so that they can use the information it yields to make informed instructional decisions. At the same time, districts and schools must develop systematic, coherent, and reliable assessment programs that ensure consistency within and across grades while complementing and building on informal assessment efforts already underway. This Knowledge Brief explains the importance of early assessment and identifies some of its basic purposes; describes the challenges of assessing young children; explains some basic approaches to literacy assessment and how they align to specific purposes; and identifies some of the issues that need to be addressed if schools are to undertake valid and reliable literacy assessment whose results can help teachers better support all young readers. The brief is intended to help district administrators, principals, and other instructional leaders begin laying the groundwork for more consistent and effective use of reading assessment in the early primary grades. It is also intended to help them better understand the nuances and limitations of various instruments, including what decisions they can support. (Contains 2 figures, 25 endnotes, and 29 references.) (NKA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

CS

HAT CAT
SEE THE
HELLO MY
HOUSE HON
RUN TO TH
U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Illustration by James Yang

ED 470 748

UNDERSTANDING

Young Readers

The Role of Early Literacy Assessment

Written by
Stanley Rabinowitz,
Judy Wong, and
Nikola Filby

As educators and policymakers search for sound ways to support low-achieving schools and students, many see improved literacy as a key because, to a greater degree than with any other basic skill, the inability to read at grade level impedes student success in all other academic areas. As is often noted, students must learn to read *in primary grades so they can* read to learn *in later grades.*[1] *This belief is a centerpiece of President Bush's "No Child Left Behind" education plan, which places priority on reading achievement in early grades. But how to improve early literacy? Effective assessment can point the way.*

WestEd
Improving education through research, development, and service

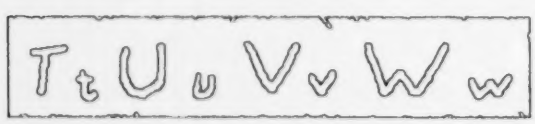
BEST COPY AVAILABLE

CS 511 508

With reading, as with all academic content and skills, effective instruction is informed by sound assessment. A lot of solid professional judgment and good informal literacy assessment now guides instruction in many classrooms. However, there is increasing recognition that if we are to ensure every child's reading success, we cannot depend solely on the idiosyncratic nature of teachers' reading assessment skills and inclinations. We must continue to develop teachers' knowledge about classroom-embedded reading assessment so they can use the information it yields to make informed instructional decisions. At the same time, districts and schools must develop systematic, coherent, and reliable assessment programs that ensure consistency within and across grades while complementing and building on informal assessment efforts already underway.

This Knowledge Brief explains the importance of early assessment and identifies some of its basic purposes; describes the challenges of assessing young children; explains some basic approaches to literacy assessment and how they align to specific purposes; and identifies some of the issues that need to be addressed if schools are to undertake valid and reliable literacy assessment whose results can help teachers better support all young readers.

The brief is intended to help district administrators, principals, and other instructional leaders begin laying the groundwork for more consistent and effective use of reading assessment in the early primary grades. We also hope to help them better understand the nuances and limitations of various instruments, including what decisions they can support.[2]



The Roles of Reading Assessment

The value of establishing a firm reading foundation in a child's early school years is incontrovertible. Research shows that students not reading at grade level by the end of the third grade are more likely than their peers who are reading well to experience difficulty reading throughout their school career, to perform poorly in other subjects, and to eventually drop out.[3,4] Research also shows that the last half of the kindergarten year and the first half of first grade are optimal times for identifying reading problems; if initial reading evaluation is delayed beyond this ideal time period, it may become harder to correct serious difficulties for many children at risk of academic failure.[5] The goal, then, is to identify and begin supporting these readers as soon as possible, using appropriate assessment instruments.

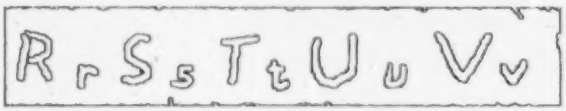
Efforts to improve student reading must be based on knowledge about what individual students and groups of students can and cannot do as it relates to literacy. Thus, effective assessment, based on specific criteria, is essential. This requires choosing the right approach and, as needed, the appropriate tool for the intended purpose. A wide range of techniques and tools are available for getting a handle on student reading in the early grades, which, for this brief, covers kindergarten through grade 2. In deciding which approach to take, it helps to start by identifying intent.

Assessments may be given for a variety of sometimes overlapping reasons,[6] chief among them: *accountability* (i.e., gathering information about student performance in order to hold students accountable for what they're expected to learn and hold districts, schools, and teachers accountable for what they're expected to

teach); *program evaluation* (i.e., understanding the degree to which a state or local reading program is effective across the broad range of students); *monitoring growth and achievement* (i.e., understanding how each child's reading is developing relative to agreed-upon standards, benchmarks, and personal goals); *screening* (i.e., identifying children at risk of having reading problems); and *diagnosis* (i.e., understanding why a child's reading is not progressing as expected, and identifying what his or her specific learning needs might be in order to instruct accordingly). No single assessment approach or tool can reasonably serve all purposes.

The range of instruments available to measure different aspects of student reading can be classified in a number of ways, and some instruments may fall into more than one category depending on how they are used. To avoid some of this confusion, this brief considers reading assessment largely in terms of purpose. To that end, it draws a general, although not rigid, distinction between two categories of standardized assessments:[7] large-scale, group-administered assessments and those that are individually administered. Depending on their specifics and the age of the students being tested, assessments in the former group may be useful for school accountability purposes and for program evaluation. They can also be useful for gauging relative student achievement and, in doing so, for screening; that is, identifying children who appear to be lagging behind and about whose reading a teacher may need more specific and in-depth information.

For a variety of reasons, most large-scale, group-administered assessments are *not* useful for a specific diagnosis of students' learning needs. That role is more effectively carried out by individually administered assessments (sometimes in combination with group-administered tasks). Yet, while individually administered assessments typically yield much richer information about student learning than do group-administered tests, they are generally not the first choice for accountability or program evaluation because, compared to group-administered assessments, they are time-consuming, expensive, and more subject to variation in administration, scoring, and interpretation.



Standardized, Group-Administered Assessments

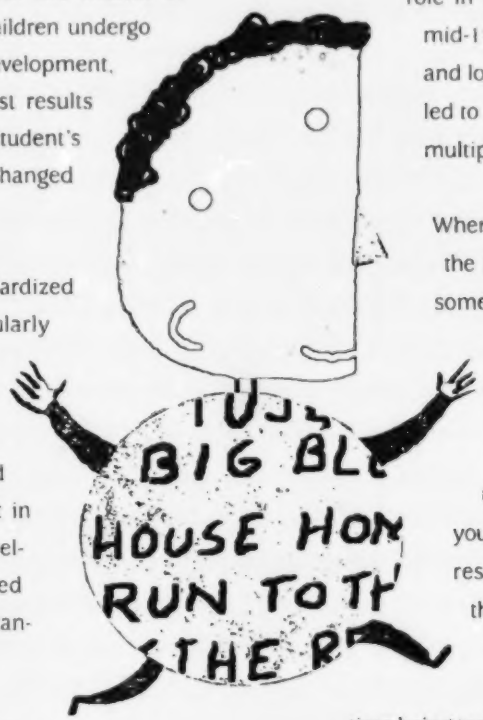
As part of the ongoing effort to ramp up achievement for all students, many states have adopted education accountability systems that incorporate large-scale assessment. State systems vary in the grade levels at which they require testing to begin, but virtually all the major commercial standardized test series offer assessments for K-2 students. So depending on the state in which they are teaching, classroom educators in the early primary grades may well be expected to administer standardized tests to their young students. Compared to tests administered in later grades, these early grade assessments typically differ in several ways aimed at making them more appropriate for younger test-takers. For example, the proctor may take a more active role in guiding the testing session, and students may respond directly in the testing booklet rather than having to mark a separate response form.

Such accommodations notwithstanding, some researchers question the appropriateness of assessing students in any formal, group-administered fashion prior to the third grade. Underlying this skepticism are concerns about the reliability and validity of such tests in the early grades, as well as the interpretability of their results.

Measurement specialists acknowledge that the younger the child being assessed, the greater the likelihood of measurement errors. The National Education Goals Panel noted that when used with children under the age of eight, group-administered tests have limited

validity and reliability.[8] Given such issues, some researchers suggest that using standardized tests for children under the age of eight for placement or school retention decisions may be detrimental to their ultimate achievement.[9]

Anyone hoping to use such test results for instructional purposes must also consider the matter of timeliness. Because young children undergo rapid rates of growth and development, by the time standardized test results are scored and reported, a student's skills and needs may have changed significantly.



The use of large-scale standardized assessments can be particularly troublesome for children from culturally and linguistically diverse backgrounds. These children may have had less exposure to the context in which many nationally developed test items are couched and less experience with standardized testing situations.

These and other factors[10] make it difficult to interpret young students' test outcomes, and inaccurate interpretation can, naturally, lead to erroneous decisions about student needs. Interpreting the results of large-scale reading assessments is particularly problematic. Different kinds of standardized, group-administered reading tests focus on different aspects of reading (e.g., phonemic awareness, letter knowledge, phoneme-grapheme correspondence, phonological awareness, vocabulary). However, the tasks that are typically included — such as sound and word recognition and short reading passages — tend to be unconnected to each other and require very little actual reading of continuous text. Furthermore, they invariably involve a selected-response (i.e., multiple choice) format, intended to allow efficient, computer-assisted scoring. Few of these tests require teachers to observe how a student actually reads.

Given the limitations of multiple-choice testing in general, and especially for young children, many educators have hoped that authentic or performance-based assessments would come to play a more prominent role in large-scale assessment. Yet, by the mid-1990s, concerns about the technical and logistical efficiency of large-scale tests led to a renewed reliance on standardized, multiple-choice assessments.

When such assessments are required in the early primary grades, as they are in some states for accountability purposes, accommodations should be made to lessen the stress on students. Ideally, tests should be administered by a proctor who is both familiar to the students and aware of how to make young test-takers comfortable. Frequent rest breaks should be scheduled and the test administration period should be stretched across several days so students spend a more limited time being tested each day. Prior to the testing time, students should be given ample practice on exercises similar to those found on the instrument.

No high-stakes decisions at any grade should be based solely on the results of a group-administered standardized assessment. This is particularly true in dealing with younger students. If such assessments are mandatory, their results should be used, in conjunction with other data, to help inform the overall picture of each student's and school's growth in reading achievement and of the relative strength and weaknesses of a school's reading program. Finally, any interpretation of student, class, and school results must factor in the inherent reliability and validity concerns noted above.

BEST COPY AVAILABLE

Individually Administered Assessments

In addition to any federal, state, or district decisions about collecting student reading data for purposes of accountability, program evaluation, or measuring achievement, those closest to young learners — classroom teachers, reading resource teachers, and others — need rich information about how individual students are reading. Careful assessment of individual readers will yield specific information about particular areas of a child's reading or language that need strengthening, as well as identifying existing strengths a teacher can reinforce and upon which he or she can build.

In fact, a well-trained teacher can glean much important information about individual children's reading or pre-reading skills from informal, embedded assessment. In her article, "How Do We Assess Young Children's Literacy Learning,"[11] Terry Salinger offers many examples, including that of a grocery list prepared by a kindergartner during dramatic play. The list, Salinger explains, serves as evidence that the child who created it understands some basic print concepts (e.g., text has meaning; text moves from left to right and from top to bottom of a page). The invented spelling used by the child — *krts* for carrots — reveals that the child "has a good grasp of initial, medial, and final consonant sounds and of phonemic awareness. The rendering for *bologna* is three syllables long — *bu-lo-e* — and is impressive evidence of the child's ability to segment words. ... All these pieces of assessment data are valuable in determining an appropriate instructional program for this kindergarten student."

Such ongoing embedded assessment is an invaluable source of information for a teacher, but it shouldn't be the only source. Rather, it should be a key element in a broader assessment program. Many experts have argued for an evaluative approach that is systematic,[12] "controlled, not casual,"[13] and structured or "standardized" in procedure.[14]

Now the new federal Reading First initiative, which provides states with money for reading programs to allocate to districts on a competitive basis, makes funding contingent on using formal and systematic reading assessment. Specifically, it requires that subgrantees — local education agencies — "select and administer screening, diagnostic, and classroom-based instructional reading assessments to determine which students in kindergarten through grade three are at risk of reading failure." Some states have also begun requiring diagnostic assessment. The Utah Performance Assessment System for Students (U-PASS), for instance, which had already required schools to administer criterion-referenced tests (CRTs) at each grade, recently added to its requirements the use of diagnostic reading tests in each grade, up through grade 9, for those students falling below state-developed performance levels on the CRTs. In grades 3 and up, districts can choose from among three group-administered, multiple-choice reading tests. But for grades 1 and 2, the state has approved three individually administered, constructed-response tests. Utah has also instituted a kindergarten screening test. During the first two weeks of school, each entering student is tested individually; post-testing is done in the spring. While not part of U-PASS, this assessment is also used to guide reading instruction.

Many good methods exist for getting an in-depth understanding of a child's reading strengths and weaknesses. One's choice of assessments and how the information they yield will be used are naturally influenced by one's understanding of literacy and how it develops, as well as by the resources a district or school can commit to the effort, including time and professional development. Whatever forms it takes, individual assessment of young readers and prereaders for instructional purposes should be systematic, efficient, consistent across teachers within a school, and able to yield valid and reliable results.

Assessment must also be tied to a common understanding of what students should know and be

Common Stumbling Blocks to Reading

A body of research points to weak phonemic awareness as the seed of most early reading problems.[17] Phonemic awareness doesn't affect a child's ability to speak or to understand spoken language, but those who can't easily isolate and manipulate the underlying sounds in spoken words called phonemes will naturally have a tough time learning how to match those sounds to letters and letter combinations, which is essential to decoding. Children start school with varying degrees of phonemic awareness, but with rare exception their understanding of and control over phonemes can be developed. Early identification of children's relative phonemic awareness tells a teacher how much emphasis to place on this effort, for the class as a whole, for groups of students, and for individual students.

A child's knowledge of letters — their names and sounds — is also considered highly predictive of reading problems absent intervention.[18] An incoming kindergartner who does not know the alphabet has likely had few literacy-building opportunities and, therefore, probably needs focused help.

able to do at key points in the reading development process, thereby linking assessment to instruction. Within the exceptionally broad range of "normal" development among young students, there must nonetheless be agreed-upon standards and benchmarks so that teachers, parents or guardians, and, when appropriate, students, themselves, know how they are progressing in light of those expectations. One example of such a system is "A Framework for Early Literacy Instruction"[15] developed by Mid-continent Research for Education and Learning. In addition to including components aimed at ongoing diagnosis of students' learning needs, an effective assessment program will also include a component designed to anticipate any reading problems so instruction can be planned to address them before a child starts falling behind.

ASSESSING AS A PREVENTATIVE MEASURE

What appears clear from research and practice is that while a certain proportion of students will become successful readers almost irrespective of the instruction they receive, others will require instruction targeted to their needs if they are to succeed.[16] For that reason, many researchers now speak of assessment in the context of prevention, the goal being to identify students in this latter group and plan instruction to address their learning needs. Careful early assessment can help identify them and help teachers begin to understand what kind of instruction — including its intensity and duration — might best serve individual students or groups of students. In this sense, early identification may refer to a time early in a child's formal schooling (i.e., in kindergarten) or early in each school year. Research has identified some specific predictors for reading problems, and a number of early-grade assessment instruments focus on these predictors. (See "Common Stumbling Blocks to Reading.")

Some argue that assessing for early identification of students who would most profit from focused attention in reading is most effective — and efficient — when started no sooner than midyear in kindergarten. Torgeson,

for example, argues that "given the widely varying range of children's preschool learning opportunities, many children may score low on early identification instruments in the first semester of kindergarten simply because they have not had the opportunity to learn the skills."[19] By mid kindergarten, however, some of that variation will have been mitigated by instruction, which means assessment at this stage should more accurately predict a child's potential for reading problems.

THE USE OF SCREENING

Because diagnostic assessment is, for the most part, individually administered, it is, by nature, more time-consuming than group-administered assessment. And because the teacher focuses on one child, it calls for skilled classroom management and instructional planning to ensure that the rest of the teacher's students are productively engaged in learning activities. Thus, some recommend reserving the resource-intensive effort of diagnostic assessment largely for students deemed in need of closer attention, as identified by either a formal screening test or simply by a teacher's careful observation.

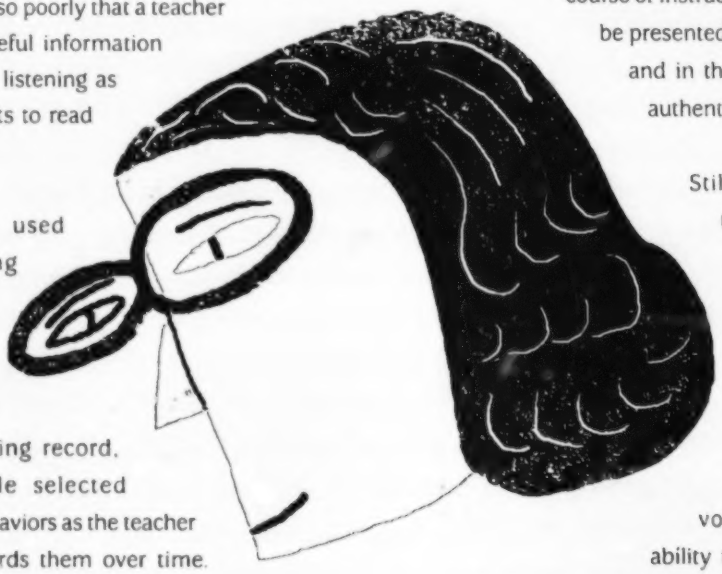
Screening instruments are intended to start the reading assessment process most efficiently by gathering some key information about each student without going into the depth — and requiring the time — of a fuller diagnostic assessment. The intent is that screening will identify those students needing specific instruction and about whom the teacher will likely need additional information, gleaned through subsequent in-depth assessment. For example, if a child is able to fluently read grade-appropriate text, it would not be necessary to assess the child's letter knowledge, phonemic awareness, or knowledge of concepts of print. The Texas Primary Reading Inventory (TPRI), a K-2 assessment program, is one example of how screening is used. The TPRI includes a screening instrument for use in each grade, as well as an inventory section designed to yield more comprehensive diagnostic data for students who don't perform well on the screening. Both sections are individually administered, but the intent is that students who do well on the screening section need not undergo the full inventory section. This is not to imply that it might not be useful to have additional diagnostic information for all students; using a screening instrument is simply a way of targeting limited resources toward those most in need of them.

Authentic or Decontextualized Assessment: Must It Be One or the Other?

As with reading instruction writ large, there has been some divergence of opinion about reading assessment, especially when it comes to young students. One school of thought, focusing on the understanding that successful reading is heavily dependent on specific skills development, sees a need to directly assess students' strengths and weaknesses in individual skill areas, often through the use of tests that are made up of decontextualized tasks, such as reading a list of individual words. The other school of thought, focusing on the understanding that successful reading is much greater than the sum of individual skills, champions assessments that yield a broader picture of a child's literacy. Evaluative efforts in this latter group, referred to alternately as performance-based, authentic, or situated assessments, tend to be contextualized within the fabric of the classroom, so much a part of classroom practice that students may not even realize they are being assessed. Such assessments also tend to be part of an ongoing evaluative process rather than periodic, extraordinary events.

Among the salient characteristics of authentic reading measures are: framing the reading assessment in an interactive, constructive process; using longer reading samples taken from sources originally written

for students to enjoy and understand rather than passages written specifically for testing; and asking students to respond to open-ended questions that allow for more than one interpretation and a range of acceptable responses.[20] Obviously, authentic reading assessment will look different in the first grade than in the sixth grade, but the same general philosophy applies: Ask children to do something real rather than contrived or decontextualized so their performance will likely be more natural and, therefore, yield more valid results. That said, an authentic assessment will not necessarily assist a teacher who seeks guidance in helping a struggling young reader. For example, some children read so poorly that a teacher may glean little useful information from watching and listening as the student attempts to read continuous text.



A commonly used authentic reading assessment tool is the error-recording-and-analysis protocol known as a running record, used to chronicle selected student reading behaviors as the teacher observes and records them over time. While many teachers observe children reading and make mental note of what they see, a running record serves as a tool for standardizing both the process of observation itself (so the teacher — and his or her colleagues — are always looking for particular things) and the recording and analysis of that information. When making a running record, the teacher should be listening to the child read whatever he or she would naturally be reading anyway in the normal course of that teacher's instruction. Because the teacher is listening and watching with specific reading behaviors in mind (e.g., the child's self-correction facility), running records made over the course of the year provide insight into a reader's progress and his or her strengths and weaknesses.[21]

Further along the spectrum from authentic to decontextualized assessment is an informal reading inventory (IRI), one of the most common methods for assessing reading comprehension found in published assessments. In an IRI, children are first asked to read a passage of text that has been leveled to age or grade and then asked to answer some questions about the text, to retell the story in their own words, or to make some inferences from what they read. In published tests, the passages have generally been written explicitly for the test, so in that sense they are less authentic than what a child is asked to read in the normal course of instruction. IRIs also tend to be presented as on-demand tasks and in this, too, they are less authentic.

Still further yet along that spectrum are direct skills tests requiring students to read lists of disassociated words as a means of checking either their sight word vocabulary or their ability to decode unfamiliar words. Finally, at the farthest end of the spectrum are those assessments that ask children to decipher nonsense words. Proponents suggest that there is no more valid measure of essential reading skills, which can otherwise be assessed only indirectly. Moreover, they note that, assuming the assessment is valid, direct skills assessment would seem to offer the advantage of more easily achieved reliability and greater ability to pinpoint strengths and weaknesses. But skeptics suggest that children don't necessarily reveal their true abilities when undertaking tasks that have no meaning to them.

While some might argue for using assessments from only one side or the other of this continuum, Torgeson

suggests that these assessment approaches should, instead, be considered complementary: "The goal of 'authentic assessment' is to measure children's application of broad literacy skills to authentic tasks, like gathering information for a report, or ability to read a selection and then write a response to it. It also seeks to measure children's enjoyment, ownership, and involvement in literacy skills at school and at home.... All of the literacy outcomes that are part of authentic assessment are important parts of a total literacy assessment program....

"However," Torgeson continues, "since these procedures are focused on high-level reading outcomes, they cannot provide precise information about level of performance in important subskills of reading. If a child's overall performance on authentic literacy tasks is limited, it is frequently difficult to obtain a precise estimate of the specific component skills that are weak." The goal of skills assessment, he concludes, "is to quantify the degree of skill a child possesses in word identification processes that have been shown to be a critical foundation for overall reading success."[22]

WHAT A LOCAL READING ASSESSMENT PROGRAM CAN LOOK LIKE

When developing a local assessment program, it helps to understand the range of published assessments available, some common early-grade assessment tasks and when they are generally administered, and how others have blended various assessment approaches into a systematic program. To find out what published assessments are already available, see the reading assessment database for kindergarten through grade 2 developed by the Southwest Educational Development Laboratory[23] and *A Practical Guide to Reading Assessments* (Kame'enui, Simmons, & Cornachione). As for what to assess when, Figure 1 presents some common K-3 assessment tasks along with the grade levels at which they are usually assessed. What follows are short descriptions of several assessment approaches that incorporate individually administered assessments.

An Observation Survey of Early Literacy Achievement. The observation survey (Clay, 1993) is an integrated assessment of students' emerging literacy, allowing teachers to monitor progress and, ultimately, identify children in need of more targeted instruction. In addition to the use of a running record, the system includes five subtests used to observe and analyze a child's understanding and use of literate behaviors. The tests include 1) letter identification, 2) Ohio Word Test, 3) concepts about print, 4) writing vocabulary, and 5) dictation. In its use of a running record, the observational survey adheres to principles of authentic assessment, whereas its subtests embody the kind of direct skills assessment Torgeson recommends for children who appear to be struggling with their reading efforts.

Texas Primary Reading Inventory (TPRI). The TPRI contains an individually administered screening and inventory section for each grade. Kindergarten assessment is done at both midyear and the end of the year; in grades 1 through 3, administration is recommended at both the beginning and the end of year, but may be given midyear as well. In each grade, the screening section includes relatively few tasks. At the beginning of first grade, for example, the screening starts with children asked to identify the names and sounds of 10 letters. Those who get 8 or more correct are then asked to read a list of 8 words, while those who get fewer than 3 correct are given a phonemic awareness task. For those who've moved to the word list task and get 3 or more correct, the screening stops; the teacher then has them do the reading comprehension task from the inventory.

For the phonemic awareness screening task, children are asked to blend 6 different sets of phonemes. Those who get 5 or more right move to the comprehension task of the inventory, while those who get fewer than 5 correct move, instead, to the more in-depth inventory tasks. Depending on how they score on individual inventory tasks, students could be asked to do as many as 13 tasks, covering a range of elements and

skills, including book and print awareness, rhyming and blending word parts, medial vowel substitution, and reading comprehension. For children who are unable to read the story in the comprehension task, the teacher reads it aloud and the task becomes one that assesses *language* comprehension rather than *reading* comprehension.

Literacy Portfolios. A portfolio is a purposeful collection of student work that exhibits to the student or others his or her efforts and achievements in one or more areas. Not all collections of work are portfolios. To be considered a portfolio, the tasks must be chosen intentionally and must be complementary so that, when taken together, they provide a comprehensive measure of students' reading progress and overall achievement.

Portfolios can be used to involve not just the student and teacher in making decisions about supporting student learning, but also parents/guardians and peers, who can contribute significantly to student learning in a variety of ways. Portfolios allow for the widest array of assessment forms. That said, if a portfolio-based assessment system is to be coherent, teachers within and across grades must carefully plan what to include for what purpose in what grades. Moreover, in evaluating student work samples, teachers must agree on performance standards and benchmarks.

A critical issue in the use of portfolios is the credibility of the various portfolio entries and the value of portfolio evidence outside the classroom.[24] Do the entries reflect authentic learning, applicable beyond

Figure 1: Common K-3 Assessments

Element of Reading	Common Time to Assess				Typical Methods of Assessment
	K	1	2	3	
Phonemic Awareness: Ability to notice and work with the individual sounds in spoken words	◔				Sound matching, segmenting, or blending
Letter-Sound Correspondence: Understanding the alphabetic principle — the relationship between written letters and spoken sounds		◔			Pronouncing words or finding the word a teacher pronounces; running records; writing samples
Fluency: Ability to read text quickly and smoothly		◔	◔	◔	Rate of reading; phrasing and expression
Vocabulary: Knowing what words mean	◔	◔	◔	◔	Classroom assessment; oral activities
Comprehension: Understanding oral language and written text; monitoring meaning	◔	◔	◔	◔	Story retelling; questions and answers
Attitude and Interests: Motivation to read and attraction to different topics, authors, texts			◔	◔	Observation; interviews

Based on: Armbruster, Lehr, & Osborn (2001); Brauger & Lewis (2001); Paris, Paris, & Carpenter (2001); Wren (2002); several state assessments.

Elements of Reading are from ***Putting Reading First*** (Armbruster), with the addition of Attitude and Interests.

the specific activity documented in the portfolio? Practitioners also worry whether the rich, qualitative evidence provided by portfolios can be fashioned to meet accountability concerns without trivializing, distorting, or undercutting the instructional value of such evidence. Finally, the use of portfolios may create burdens for teachers, students, and schools for keeping track of progress toward completion of required tasks, storage of "works-in-progress" and completed work, scoring, and reporting.

Significant amounts of planning and training are required to properly implement a portfolio assessment system. With great commitment and, in most cases, the help of an outside consultant, schools can develop their own literacy portfolio systems.

Michigan Literacy Progress Profile (MLPP) and Portfolio. The MLPP provides a comprehensive assessment plan for preschool to third-grade students. Aligned to the state's English language arts standards, the system includes two types of assessments, a framework from which to make assessment decisions, and, related to each assessment, suggested interventions and strategies that can be used in the classroom and at home. Because its overall purpose is to provide diagnostic data for guiding instruction, the assessment is not normed. Its intent is to help teachers understand what each child can do so teachers can help every student move forward irrespective of their individual starting points.[25]

The first set of assessments is designed to identify children's development in "milestone" behaviors related to reading: oral language, fluency, comprehension, writing, and attitudes and self-perception. The Michigan Department of Education, which developed the MLPP with a committee of teachers, parents, and teacher educators, recommends that milestone behaviors be assessed three to four times a year. Depending on a child's performance on these initial assessments, teachers may use more targeted assessments to "dig deeper" in understanding the student's discrete "enabling skills" (e.g., concepts of print, letter/sound identification, phonemic awareness, sight word/decodable word identification).

The system includes a variety of rubrics for scoring purposes. There is, for example, a retelling rubric for grades K-2 based on narrative text that students have read or listened to. The rubric identifies and describes four performance levels for each of four "qualities of retelling." (See Figure 2.) The level of a student's retelling performance is considered in relationship to the level of the text. As the text level rises, the performance on some of the retelling qualities may drop, but should rise again as the child's literacy skills further develop. A student's performance on specific elements of retelling should guide the teacher's subsequent instruction.

Students' literacy profiles present a picture of their developing literacy, as evidenced in assessment scores and documented teacher observations over time. Michigan's Department of Education also recommends that, if it fits with a school's assessment philosophy, educators include the profile as part of a larger reading-and-writing portfolio for each student. The portfolio would also include student-chosen work samples with the student's reflections on the work, along with work artifacts chosen by the teacher to illustrate the student's literacy at the beginning, middle, and end of the year. The intention is for the portfolio to travel with the student as he or she moves up the grades and/or changes schools.

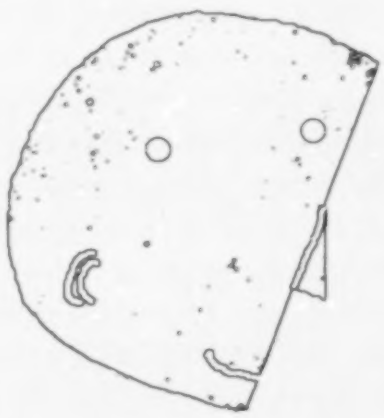


Figure 2: MLPP 2001 Retelling Rubric Grades K-2 Narrative Text

Qualities of Retelling	4–Mature	3–Capable	2–Developing	1–Beginning
Gist/Main Idea • Lessons Learned • Plot Main Idea	Retelling includes a clear generalization that states or implies the plot/main idea and lesson learned.	Retelling includes a generalization that states or implies the plot/main idea from the story.	Retelling indicates inaccurate or incomplete understanding of the plot/main idea.	Retelling includes minimal or no reference to, or understanding of, plot/main idea.
Story Elements Organization Linguistic Spillover	Retelling contains a clear statement of all story elements (main characters, setting, problem, major events, and resolution) and their connection to one another.	Retelling contains a clear restatement of most story elements (main characters, setting, problem, major events, and resolution) and their connection to one another.	Retelling contains a restatement of some story elements with minimal connections to one another.	Retelling contains minimal restatement of story elements.
	Events are retold following a logical sequence with a beginning, a middle, and an end.	Events are retold mostly in appropriate order with a beginning, a middle, and an end.	Events are retold in a somewhat disconnected fashion. The beginning or middle or end may be deleted.	Events lack sequence.
	Use of language, conventions, and/or format from the selection reflects an elaborated and personalized understanding of the story.	Use of language, conventions, and/or format from the selection indicates basic understanding of the story.	Use of language, conventions, and/or format from the selection may indicate superficial understanding.	Retelling includes little or no use of language, conventions, and/or format from the story.

Included for review purposes only with permission from the Michigan Department of Education, Office of School Excellence, Curriculum Leadership.

Conclusion

For instructional planning purposes and to prevent and/or identify delays in children's reading development, ongoing, individually administered performance assessment provides the most valid information. The National Research Council notes that a single test should not be the sole basis for high-stakes decisions about individual children, such as those relating to promotion or the need for special services. Other factors, such as teacher observation, should also be considered. Early childhood professionals support the need for assessing K-2 students by using multiple measures over time and in the context of daily classroom activities.

Most specialists believe that no high-stakes accountability testing of individual readers should be done before the end of the third grade. Instead, assessment programs should be designed to gather data for screening and diagnostic purposes, as well as to understand the conditions of learning. For these purposes, current research and expert opinion suggest using a range of short direct-assessment tasks that include specific knowledge/skill measures, in addition to a variety of more complex and authentic tasks that, minimally, entail reading connected text. In addition to finding out how a child is achieving in reading, as evidenced in both cognitive and affective indicators, teachers should also make efforts to understand the

degree and type of literacy support in the child's home. All this information can guide instruction, including prevention and intervention efforts.

When properly developed and administered, primary reading assessment supports the ongoing development of readers across the developmental spectrum, as well as opening the way for prevention of reading problems or for early intervention with those children in need of additional reading support. Good assessment is a key element in the effort to ensure that all children become successful readers and, subsequently, high achievers in other academic pursuits. Ongoing research is essential to identify what assessment tools are most effective across settings, student populations, and purposes.

Local decisions need to be made about the mix of primary reading assessments. Assessment should be an integral part of a reading program, supporting teachers in identifying reading problems and then knowing what to do about them. An assessment program should be appropriate to the local context. Resources must be allocated not only for the assessment program itself, but also for training teachers and other staff in the analysis of data and in how to appropriately respond to the story told by that data. Depending on the child, an appropriate response could range from continuing the status quo to changing instructional pacing or approach, to having the child work with a reading specialist outside the classroom or partake in some other intervention. In some instances, assessment data across students could prompt rethinking of the whole reading program. A key decision in this effort relates to benchmarks, or expected performance levels at specific points in time, that can guide the pacing of instruction and identify students who are falling behind and need extra help. An assessment program based on these various considerations then becomes a critical tool in any thoughtful approach to reading instruction.

ENDNOTES

[1] Gaining a solid literacy foundation in the early primary grades is critical to a child's academic success and is, therefore, the focus of this brief. This is not meant to imply that a strong foundation in the primary grades is by itself sufficient for later success. If students are to be able to successfully read academic texts in specific content areas as they move up through the grades, teachers must ensure that active reading development continues beyond third grade. See, for example, Schoenbach et al. (1999).

[2] This brief does not attempt to address the specific and unique issues related to reading assessment for special needs students (e.g., English language learners, hearing impaired students, those with phonological processing difficulties).

[3] In testimony to Congress, Reid Lyon (2001) summarized research supported by the National Institute of Child Health and Human Development and the Department of Education's Office of Educational Research and Improvement as showing that "failure to develop basic reading skills by age 9 predicts a lifetime of illiteracy. ... On the other hand, the early identification of children at risk for reading failure coupled with the provision of comprehensive early reading interventions can reduce the percentage of children reading below the basic level in the fourth grade (e.g., 38 percent) to six percent or less."

[4] Otuya & Krupka, S. (1999).

[5] Snow et al. (1998); Stanovich (1986).

[6] Katz (1997).

[7] A standardized test — as defined by the U.S. Congressional Office of Technology Assessment — is one that uses uniform procedures for administration and scoring to assure that the results from different people are comparable.

[8] Shepard et al. (1998).

[9] Stevens & Debord (2001, Spring).

[10] Shepard et al. (1998, February), p. 4.

[11] Salinger (1998b), p. 228.

[12] Fountas & Pinnell (1996).

[13] Clay (1985), p. 1.

[14] Salinger (1998b).

[15] Available at http://www.mcrel.org/resources/literacy/ela/framework.html.

[16] Diamond & Mandel (1996); National Reading Panel (2000); Snow et al. (1998).

[17] Yopp and Adams, as cited in Diamond & Mandel (1996), p. 3; National Reading Panel (2001).

[18] Ehri and McCormick, as cited in California Reading and Literature Project (2000); Torgeson (1998).

[19] Torgeson (1998), p. 4.

[20] Available at http://www.eduplace.com/rdg/res/literacy/assess2.html.

[21] New Zealand researcher Marie Clay (1993, p. 7) notes that running records have both face and content validity. "You cannot get closer to the valid measure of oral reading than to be able to say the child can read the book you want him to be reading at this or that level with this or that kind of processing behavior. Little or nothing is inferred. You can count the number of correct words to get an accuracy score. The record does not give a measure of comprehension, but you can tell from the child's response to the story and from the analysis of error and self-correction behavior how well the child works for meaning. And you can gauge his understanding of the story in the discussion you have with him about the story. You do not get a score on letters known, but you can see whether the child uses letter knowledge on the run in his reading." Wren (2002, p. 2) suggests that teachers should take care when trying to use the running record to assess both reading accuracy (i.e., word recognition) and reading comprehension. "Reading comprehension often suffers when [children] are asked to read a passage of text out loud. When children read orally, they usually concentrate on reading accurately and do not pay as much attention to comprehension of the content. Oral reading accuracy does give insight into decoding skills and strategies, but that is a separate test. A reading comprehension test is most accurate if the child is not reading aloud for an audience."

[22] Torgeson (1998), p. 7.

[23] Available at: www.sedl.org/reading/rad.

[24] Salinger (1998a), p. 188.

[25] "Reading Plan for Michigan" (1999), State of Michigan Department of Education. Also, B. Rockafellow (personal communication, May 10, 2002). For more information about the Michigan early reading assessment system, contact: Bonnie Rockafellow at rockafellowb@michigan.gov or write c/o Office of Professional Preparation, Michigan Department of Education, 608 W Allegan, P.O. Box 30008, Lansing, MI 48909.

REFERENCES

Armbruster, B., Lehr, F., & Osborn, J. (2001). *Put reading first: The research building blocks for teaching children to read (Kindergarten through Grade 3)*. Washington, DC: The Partnership for Reading: National Institute for Literacy, National Institute of Child Health and Human Development, and U.S. Department of Education.

Braunger, J., & Lewis, J. (1997). *Building a knowledge base in reading*. Portland OR: Northwest Regional Educational Laboratory, International Reading Association, & National Council of Teachers of English.

Clay, M. M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.

Clay, M. M. (1985). *The early detection of reading difficulties* (3rd ed.). Auckland, New Zealand: Heinemann Education (division of Octopus Publishing).

Denton, D. R. (1999, September). *Reading reform in the SREB states: Early assessment*. Atlanta: Southern Regional Education Board. Available online at: www.sreb.org.

Diamond, L., & Mandel, S. (1996.) *Building a powerful reading program: From research to practice*. Emeryville, CA: Consortium on Reading Excellence.

Falk, B. (1998). How the Primary Language Record supports teaching and learning. In C. Harrison & T. Salinger (Eds.), *Assessing reading 1: Theory and practice*. New York: Routledge.

Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.

Kagan S. L. (1999). Redefining 21st century early care and education. *Young Children Journal*, 54(6), 2-3.

Kame'enui, E., Simmons, D., & Cornachione, C. (n.d.). *A practical guide to reading assessments*. Washington, DC: U.S. Department of Education.

Katz, L. G. (1997, April). *A developmental approach to assessment of young children*. Champaign, IL: University of Illinois, Children's Research Center. (ERIC Document Reproduction Service No. EDO-PS-97-18)

Lyon, G. R. (2001, March 8). *Measuring success: Using assessments and accountability to raise student achievement*. Statement to the Subcommittee on Education Reform, Committee on Education and the Workforce, U.S. House of Representatives.

Michigan State Department of Education. (1999). *Reading plan for Michigan*. Available: http://goal.ncrel.org/litweb/docs/michigan.htm.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Rockville, MD: NICHD Clearinghouse. Available: www.nichd.nih.gov/publications/nrp/smallbook.htm.

Otuya, E., & Krupka, S. (1999). *Federal and state strategies to support early reading achievement*. Available online from the ETS State and Federal Relations Office: www.ets.org.

Paris, S. G., Paris, A. H., & Carpenter, R. D. (2001). *Effective practices for assessing young readers*. Ann Arbor, MI: Center for the Improvement of Early Reading Achievement, University of Michigan.

Rabinowitz, S., & Ananda, S. (2001). *Balancing local assessment with statewide testing: Building a program that meets student needs*. San Francisco: WestEd.

Salinger, T. (1998a). Consequential validity of an early literacy portfolio. In C. Harrison & T. Salinger (Eds.), *Assessing reading*. New York: Routledge.

Salinger, T. (1998b). How do we assess young children's literacy learning? In S. B. Neuman & K. A. Roskos (Eds.), *Children achieving: Best practices in early literacy*. Newark, DE: International Reading Association.

Schoenbach, R., Greenleaf, C., Cziki, C., & Hurwitz, L. (1999). *Reading for understanding: A guide to improving reading in middle and high school classrooms*. San Francisco: Jossey-Bass.

Shepard L., Kagan, S., & Wurtz, E. (Eds.). (1998). *Principles and recommendations for early childhood assessment*. Washington, DC: National Education Goals Panel.

Snow, C. E., Burns, S. M., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stevens, G. G., & DeBord, K. (2001, Spring). Issues of assessment in testing children under age eight. *The Forum, 6*, 2.

Teale, W. H. (1990). The promise and challenge of informal assessment in early literacy. In *Assessment for instruction in early literacy*. New York: Prentice Hall.

Torgeson, J. (1998, Spring/Summer). Catch them before they fall: Identification and assessment to prevent reading failure in young children. *American Educator*. Washington, DC: American Federation of Teachers.

U.S. Department of Education, National Center for Education Statistics (2001, June). *The condition of education, 2001* (NCES 2001-072). Washington, DC: U.S. Government Printing Office.

Wren, S. (2002). *Topics in early reading coherence: Methods of assessing cognitive aspects of early reading development*. Austin, TX: Southwest Educational Development Laboratory.

Wren, S. (2001). *Topics in early reading coherence: Phive phones of reading*. Austin, TX: Southwest Educational Development Laboratory. Available online only: www.sedl.org/reading/topics/phon.html.

Wren, S. (2000). *The cognitive foundations of learning to read: A framework*. Austin, TX: Southwest Educational Development Laboratory.

WestEd, a nonprofit research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. While WestEd serves the states of Arizona, California, Nevada, and Utah as one of the nation's Regional Educational Laboratories, our agency's work extends throughout the United States and abroad. It has 16 offices nationwide, from Washington and Boston to Arizona, Southern California, and its headquarters in San Francisco.

For more information about WestEd, visit our Web site: WestEd.org/, call 415.565.3000 or, toll-free, (1.877) 4-WestEd, or write:

WestEd
730 Harrison Street
San Francisco, CA 94107-1242

This Knowledge Brief was produced in whole or in part with funds from the Office of Educational Research and Improvement, U.S. Department of Education, under contract #ED-01-CO-0012. Its contents do not necessarily reflect the views or policies of the Department of Education.

© 2002 WestEd. All rights reserved.

Visit our Web site at WestEd.org

WestEd.

730 Harrison Street
San Francisco
California 94107-1242

Address service requested

Non-Profit
U.S. Postage
P A I D
Los Alamitos, CA
90720
Permit No. 87

BEST COPY AVAILABLE

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").